AD NUMBER
AD246349
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; AUG 1960. Other requests shall be referred to Wright Air Development Center, Wright-Patterson AFB, OH 45433.**
AUTHORITY
ASD notice 30 Oct 1979

WADC TECHNICAL REPORT 58-673

EFFECTS OF VARIATIONS IN SPECIFICITY OF KNOWLEDGE OF RESULTS ON THE IMPROVEMENT OF A PERCEPTUAL SKILL

Theodore E. Cotterman

Behavioral Sciences Laboratory
Aerospace Medical Division

112 900

AUGUST 1960

WRIGHT AIR DEVELOPMENT DIVISION

WADC TECHNICAL REPORT 58-673

EFFECTS OF VARIATIONS IN SPECIFICITY OF KNOWLEDGE OF RESULTS ON THE IMPROVEMENT OF A PERCEPTUAL SKILL

Theodore E. Cotterman

Behavioral Sciences Laboratory
Aerospace Medical Division

AUGUST 1960

Project No. 1710
Task No. 71605

WRIGHT AIR DEVELOPMENT DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

1,000 — October 1960 — 6-325

FOREWORD

This research was begun under the former Task 71636, "Effective Use of Knowledge of Performance in Training Equipment," with the author as Task Scientist. That task was one of several under Project 7197, "Human Factors in Design of Operator Trainers," with Dr. Ross L. Morgan as Project Engineer. Work under it is being continued under Task 71605, "Human Factors in the Design of Devices for Operator Training and Evaluation," of Project 1710, "Human Factors in the Design of Training Equipment," with Dr. Marty R. Rockway as technical supervisor of both. Dr. Theodore E. Cotterman, the author, is a member of the Training Psychology Branch, Behavorial Sciences Laboratory, Aerospace Medical Division of the Wright Air Development Division.

The data were gathered and analyzed at the Ohio State University as a portion of the services provided under Contract AF 33(616)-3076, "Collection and Analysis of Research Data on Human Factors in the Design of Training Equipment." Dr. Delos D. Wickens, who was Principal Investigator, and Dr. Henry Cross provided general supervision; Mr. Jason Black and Mr. William Pearson ran the subjects; and the analysis was performed by the various members of the project staff.

ABSTRACT

Ninety male undergraduate subjects estimated individually with respect to 120 stimulus photos how many degrees a 1/4-inch arrow would have to be turned to exactly parallel an adjacent arrowheaded line drawn across a 3-1/2-inch circle. The stimuli were presented for five seconds with ten seconds between and longer rest intervals separating sets of 24. Each set contained examples of the same 24 different stimuli in random order—correct answers ranging from 11 to 44 degrees and never being duplicated. Knowledge of results given orally after each estimation ranged in specificity from simple right-wrong information to amount and direction of error information for five experimental groups. No knowledge was given a sixth control group.

Differences in mean absolute and algebraic error per stimulus among groups reflect differences in specificity of knowledge given. It is concluded that knowledge of results increases the rate and level of learning to perform an absolute judgment of spatial extent and this effect is generally greater the more specific the knowledge.

PUBLICATION REVIEW

Walter F. Grether

WALTER F. GRETHER
Technical Director, Behavioral Sciences Laboratory
Aerospace Medical Division

EFFECTS OF VARIATIONS IN SPECIFICITY OF KNOWLEDGE OF RESULTS ON THE IMPROVEMENT OF A PERCEPTUAL SKILL

INTRODUCTION

A commonly-accepted technique for effective training is to provide the trainee with knowledge of the results of his performance, or psychological feedback. Learning is generally expected to be improved when the trainee is given information about his performance beyond what is naturally available as a result of performing the task. This general principle is supported by numerous research studies on the learning of various kinds of tasks. However, through precedent and the vagaries of researcher's personal interests, some kinds of behavior have been more extensively studied than others.

Perceptual skills are among those kinds of behaviors for which the effects of knowledge of results are less well determined. True, enough has been done to warrant the conclusion by Gibson (1953) that if knowledge of results is not absolutely necessary to the improvement of a perceptual judgment, it is at least of great value. Evidence for this is found in research on the tasks of judging whether one or two points are contacting the skin (Solomons, 1897); grading handwriting (Gilliland, 1925); estimating length (Thorndike, 1927); estimating auditory number (Taubman, 1944); estimating visual number (Minturn and Reese, 1951); and judging visual stimuli differing in size, brightness and hue (Eriksen, 1957). However, these studies do not go beyond the validation of the general principle to explore the effects of systematic variations in the knowledge. Thus, one may ask how changes in the specificity of the information, or in the precision or accuracy of the information, or in the frequency of giving information, or in the time at which it is given influence the effectiveness of the learning experience.

Specificity, as one kind of variation in knowledge of results, may be defined as the degree to which information given the learner describes the manner in which his performance deviates from criterion performance. The effects of variations in the specificity of knowledge about a perceptual response have been investigated by Hamilton (1929) and by Waters (1933).

Hamilton studied the effect of five different incentive conditions on judgments of length. Sixty undergraduate women individually made 50 attempts on each of two days to set a flexible rod, controlling the length of a horizontal bar of light, in such a way as to make the variable bar twice as long as a standard (120 mm.) one. Beginning with the sixth attempt on the second day an equal number (10) of subjects (Ss) were given the following treatments: (1) punishment - a bell sounded after each wrong response; (2) reward - a bell sounded after each correct response; (3) guess-with-punishment - a bell sounded after each wrong response and Ss then guessed whether their adjustment was long or short; (4) told-with-punishment - a bell sounded after each wrong response and the experimenter (E) said "long" or "short;" (5) knowledge - E said "long," "short," or "correct" after each response;

and (6) a control—no bell or knowledge. Analysis of error, expressed as a percentage of average error, showed all incentive conditions superior to the control condition. Told- and guess-with-punishment groups did not differ significantly, but were superior to reward and punishment groups which also did not differ significantly. The knowledge group was inferior to all other incentive groups (significantly so for guess- and told-with-punishment groups). In general, the time required for settings was uncorrelated with error and decreased with practice. At least superficially, the results of the experiment are at variance with the common sense hypothesis that performance is directly related to the specificity of knowledge of results. But, there are several possible explanations for the knowledge group's inferior performance. First, variations in specificity were confounded with variations in the time relations because the bell was sounded immediately after the response and E's remarks followed after some delay. Second, it is possible that once the Ss were sure they had made an error, they already had sufficient information to guide future responses. Finally, the bell may have been intrinsically more reinforcing and motivating and so enhanced performance relatively more than simple knowledge (cf. Brown, 1949).

Waters reported two studies on the learning of a perceptual judgment. According to an abstract, in the first study, improvement in judging the length of cardboard strips seemingly was unrelated to degree of information given. In the second, estimations of a 12-second interval improved in proportion to degree of information. Thus, the effect of specificity may depend on the nature of the task being learned.

Experimental variations in specificity also have produced contradictory results in the learning of nonperceptual tasks. Ross (1927) had Ss in small groups individually practice making as many tallies (four vertical lines with one diagonal) as they could in a one-minute trial. Those who had the opportunity to see the previous day's paper marked with errors and scores and a frequency table for the group and who were urged to watch progress and prepare learning curves performed better than those who were simply told who was above and who was below average. Those receiving no information were worst by a slight margin. No change in effects was noted on the last two trials when those receiving the most knowledge were given none and the others were given full knowledge. However, this finding was not confirmed in several experiments Ross (1933) performed in classroom situations. Different degrees of knowledge of weekly test performance did not produce differential learning.

In a classic experiment, Trowbridge and Cason (1932) studied improvement in drawing three-inch horizontal lines while blindfolded. Four groups of 15 male and female students were differentially treated in 100 practice trials according to whether E responded after each trial by saying nothing; or saying a nonsense syllable; or saying "right" if within 1/8 inch, otherwise "wrong;" or saying "plus X," "minus X," (indicating 1/8 inch error units) or "correct." As indicated by mean percent correct and average error, those receiving information on amount and direction of error were far superior to the others. Those receiving right-wrong information were better than those receiving no information, and those given nonsense syllables

were worst of all. In a second series of 100 trials Ss in each group were divided equally among the three conditions not encountered in the first series. The same relative performances were noted for the various conditions.

Rock (1936) found that knowledge of the number of correct responses made in a 100-item series had but slight effect on the learning by children of which of the numbers 1 through 5 were associated with which of eight stimulus words presented successively in random order. Adding knowledge of number correct to knowledge of "right or wrong" after each response resulted in slight advantage. Hirsch (1952) compared the effects on retest performance of several knowledge-of-results conditions used in conjunction with test films. Presentation of a neon light after each correct response along with the question and the correct answer proved superior to either the neon light alone or the neon light along with the number of the correct answer. Simply reshowing the original training film was as helpful as the neon light with the question and correct answer, but the best condition was a combination of both. (To the extent that recall of test questions and responses is stimulated by it, the film affords considerable knowledge of results.)

Parks (1954) found that upon test Ss performed the best tracking and ranging on a flexible gunnery device if they were given their time scores for each of the six directions after each training trial. This condition was contrasted with time scores for the three dimensions after each trial, group means after training sessions, and personal conversation after each trial. The effect upon the same kind of skill of three other conditions was explored by Goldstein and Rittenhouse (1954). They compared the value of a buzzer signal when on-target simultaneously in all dimensions with a verbal statement after each trial (including proportion of time on-target and comparisons of performance on different trial segments and with the previous record) and with a tuition condition (consisting of the verbal statement plus statements about S's specific error tendencies). In general, the amount and kind of knowledge of results made little difference, although according to performance the buzzer was best, tuition next, and verbal statements last. However, after use of the buzzer was discontinued performance deteriorated to control levels and upon transfer to a slightly different task actually became worse than control performance.

In general, many of these studies do not offer critical evidence on how the specificity of knowledge of results is related to its effectiveness for learning. It is apparent that precise control of specificity was often lacking. For example, in Ross's classroom experiments Ss still had some knowledge of results through classroom discussion of frequently missed problems. Other experiments confound guidance conditions (Ross, 1927), or time of giving the knowledge (Hamilton, Hirsch, Goldstein and Rittenhouse) or different levels of specificity (Hirsch and Rock) with specificity per se in a way as to make inferences from them hazardous. However, the experiments by Trowbridge and Cason and by Parks clearly suggest that performance is directly related to the degree of specificity of the information given the trainee about his performance.

Purpose and Hypotheses. Perceptual judgments are pervasive in the work-a-day world. In some cases they constitute the primary portion of a task: in others they are necessary portions of complex tasks on which other portions depend. In any case, whenever training or a training device must be designed to aid the learning of a task involving perceptual judgments, it is important to consider whether knowledge of results must be provided and, if so, with what degree of specificity. More exact knowledge of the effects of variations in specificity would be helpful in making such design decisions - especially when the provision for highly specific knowledge or scores would entail considerable time and/or expense.

The experiment being reported was to serve the purpose of exploring the effects of variations in the specificity of knowledge of results on rate and level of learning of a simple perceptual judgment. The judgment selected for use was the absolute one of estimating the extent of angular separation to the nearest degree. Some normative data are available for judgments of this sort as a function of the physical characteristics of the stimulus (Reese, 1953; Baker and Grether, 1954). However, to the writer's knowledge, no information is available on the particular stimulus arrangement used or on the learning of this skill. The stimulus arrangement was selected because of the likelihood of encountering displays having similar characteristics in weapon systems.

Stated formally, the hypotheses under test are as follows: (1) Rate and level of learning to estimate angular separation are increased when knowledge of results is given. (2) Rate and level of learning to estimate angular separation are directly related to the specificity of the knowledge of results given.

METHOD

Subjects. The subjects (Ss) for the experiment were 90 male college students divided equally and at random into six groups. Eight of these Ss were replacements for others which were rejected. Of those rejected, five were misrun as a result of experimenter's error, one had participated in a pilot study, one lacked sufficient time to complete the experiment, and one failed to follow instructions.

Stimuli. The stimuli consisted of 5 x 6-inch, white, matte photos. Centered on each of them was a $3\frac{1}{2}$-inch circle in bold outline with an arrowheaded line running completely across it and a small ($\frac{1}{4}$-inch) arrow adjacent, but not parallel, to the line. Four fine lines radiating 1/16 inch outward from the periphery indicated main directions. The stimuli differed as to direction and position of the arrow and the line so that there were 24 different arrangements. In an equal number, or three, of these the arrowheaded line pointed at each one of the eight main compass points - passing in one through the circle's center and in the other two a perpendicular distance of 3/4 inch on either side of center. The small arrow was randomly positioned anywhere along the line (except near the circle perimeter) and

from 3 to 12 mm. away from the line at its nearest point. Although it was oriented in the same general direction as the line, the arrow deviated by 11 to 44 degrees from being parallel to the line and the amount of deviation was never duplicated.

Five copies of each of the 24 different arrangements, or 120 stimuli in all, were used in the experimental series. These were grouped into five sets of 24 each so that each set contained copies of all the arrangements but in a different random order. For convenience in handling, the stimulus-photos were cemented to thin, white, 6 x 6-inch metal squares so that the side and bottom edges coincided.

In addition, a sample stimulus was prepared for use in instructing Ss (see Figure 1). It differed from the experimental stimuli in that the arrow-headed line does not point towards a major compass point and the small arrow deviates by a combination of direction and angular amount not used in the experimental series. Illustrations of the experimental stimuli may be seen in Appendix A.

Apparatus. The presentation apparatus consisted of a vertical board (3 x 2½ feet) mounted on wooden feet and placed on a table in front of the seated S. A 5 x 7-inch aperture was cut in the board so as to be directly in S's line of sight. Behind the aperture was a stimulus holder tilted slightly so that a 40-Watt lumiline bulb mounted just above provided fairly even illumination of the stimulus. The board obstructed the S's view of the exposed upper portion of the metal squares in such a way that only the 5 x 6-inch stimulus-photos mounted on them were normally visible. A horizontally sliding panel, somewhat larger than the aperture, permitted the experimenter (E) to control the duration of S's view of the stimuli. The whole apparatus was painted dull black. A desk lamp used by E behind the apparatus constituted the main source of light in the small sound-protected cubicle in which the experiment was done. This lamp was arranged so that to S the top and side edges of the board appeared evenly-surrounded by light of moderate intensity, thus alleviating the unpleasantness of the contrast between the brightly-lighted white stimuli and the black board.

Procedure. Upon reporting for the experiment each S was seated before the apparatus and read the same general instructions. The instructions informed the S that each of a series of stimulus cards, like the sample, would be displayed through the aperture for about five seconds and requested that he respond to each card as follows:

"For each card imagine that you are flying a plane in the direction shown by this small arrow. The long line with the arrow on one end shows the direction in which you are supposed to fly. I'd like you to tell me how many degrees you would have to turn to fly in the same direction as the long line, that is, so that the small arrow would be parallel to the line."

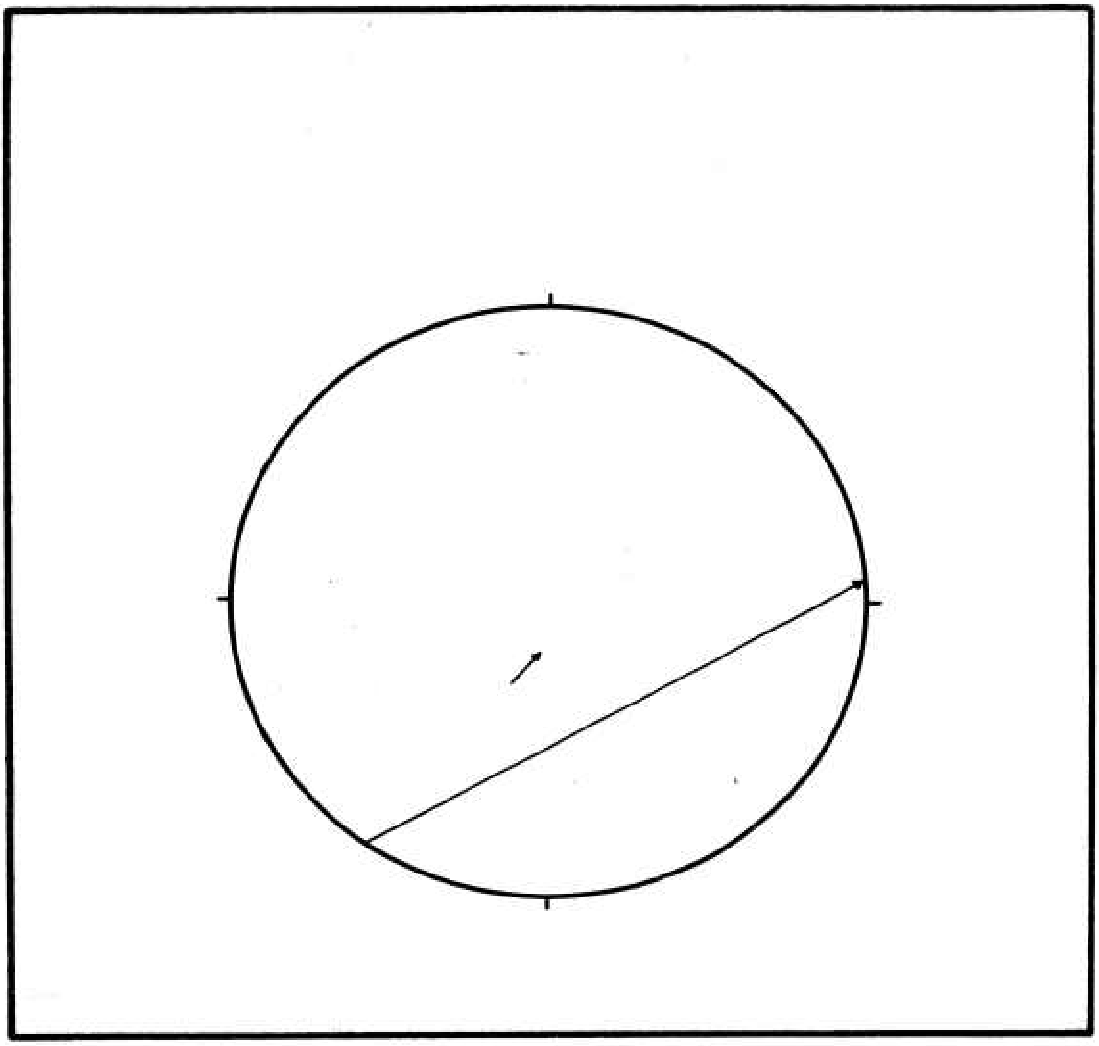


Figure 1. The Sample Stimulus (20°)

After giving the general instructions the E continued with an additional statement according to the S's assignment to one of the six treatment groups. These statements described the kind of information the S would be given about each estimate. Accordingly, information given to five of the groups ranged in specificity from a statement as to whether the estimate made was right or wrong, to a statement of the correct answer. No information was given to a sixth, control group. Table 1 contains a summary of these different treatments. A complete copy of the instructions read to Ss may be found in Appendix B.

TABLE 1

Rules for Giving Knowledge of Results to the Various Groups

Group	Knowledge of Results
I	no information
II	"right" if either correct to nearest degree or no more than one degree in error; otherwise, "wrong"
III	"right" if correct to nearest degree; otherwise, "wrong"
IV	"over," "under," or (if to nearest degree) "correct"
V	"over x degrees," "under x degrees," or (if to nearest degree) "correct"
VI	"x degrees" (the correct answer to nearest degree)

All Ss were shown successively each stimulus of each of the five 24-stimulus sets. Using a stopwatch, E presented a stimulus for approximately five seconds, allowed 10 seconds to elapse during which he noted S's response and gave the appropriate knowledge of results, and then presented the next stimulus. S was permitted a 2-minute rest between each set, with the exception that he was given a 6-minute rest between the third and fourth sets. Thus an experimental session required 42 minutes in addition to the time required for giving instructions.

The stimulus order within each of the five sets was always the same, enabling the convenient recording of responses on a prepared record sheet. However, the order of presenting the five sets to Ss within each group was in accordance with 15 quasi-random orders. The orders were restricted so that each set was first three times and last three times and all 20 possible digrams were represented at least once but no more than five times. Thus, within a group the sets were presented in a different order to each S, but the same orders were used for all groups.

Scoring and Analysis. For each presentation or stimulus, subtraction of the correct answer from the corresponding judgment produced a score indicating both direction and the amount of error. In subsequent analysis these were treated in two ways - as absolute error (with sign disregarded) and as algebraic error. The 5% level was used in all tests of the significance of differences.

RESULTS

Effects of Knowledge. Mean absolute error per stimulus for the various groups is plotted against sets of stimuli in Figure 2. The values plotted may be found in Appendix C along with mean absolute error per stimulus computed over blocks of six stimuli. The within-group fluctuations in means over six stimuli indicate that performance on this task was quite variable. Mean absolute error on the very first stimulus for Groups I through VI was computed to be 8.7, 7.5, 8.8, 8.5, 7.7, and 9.8 degrees respectively. Since these means do not differ significantly (F-ratio of 0.20) the groups are regarded equal in absolute error before differential treatment.

From Figure 2 it is evident that all except the no-knowledge group decreased in absolute error with practice. For all groups except Group I the reduction in mean error between the first and the last stimulus sets is significant according to *t*-tests summarized in Table 2. Inspection serves to indicate, however, that the groups did not all improve at the same rate. Some improved very gradually while others had virtually reached their terminal level in the third set. The best groups showed surprisingly little error after training - only about four degrees.

TABLE 2

Significance of Differences in Mean Error Between First and Last Sets of Stimuli (*t*)

Group	Absolute Error	Algebraic Error
I	1.22	1.48
II	2.79*	0.87
III	3.42**	0.54
IV	3.73**	0.85
V	2.80*	1.50
VI	5.55**	0.09

* $p < .05$ ** $p < .01$ (14 *df*)

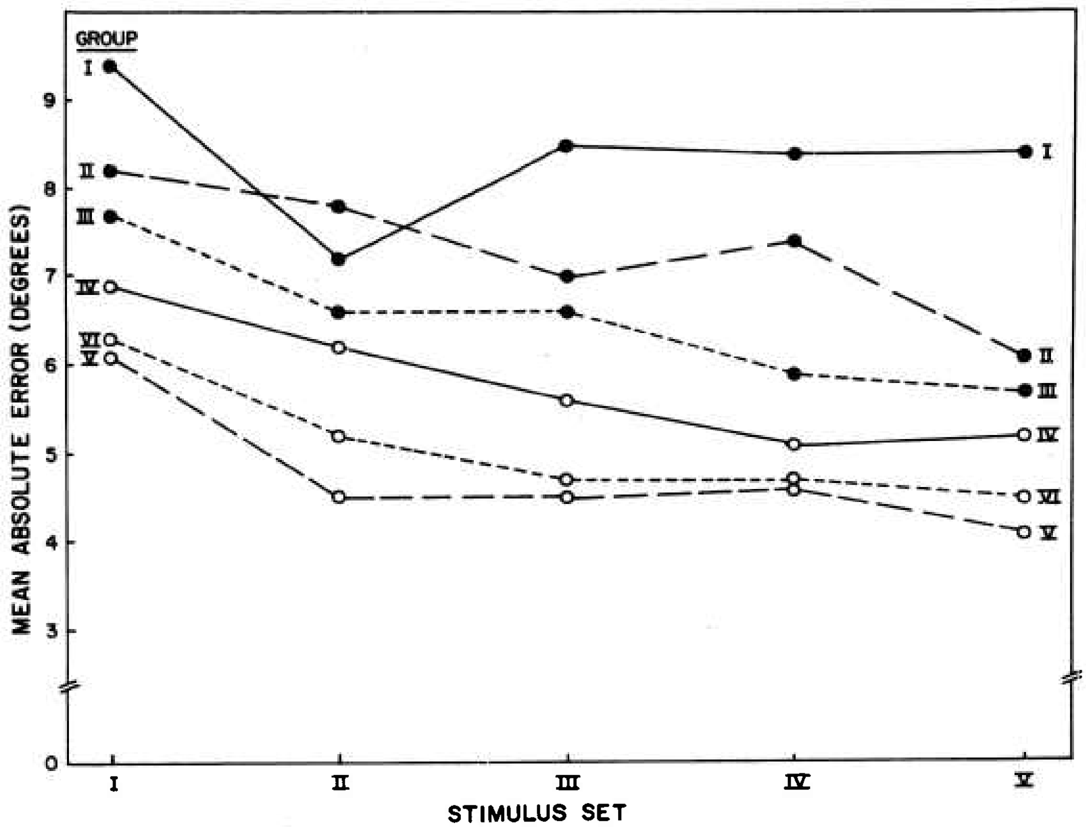


Figure 2. Mean Absolute Error Per Stimulus Plotted Against Sets of Stimuli.

It also is apparent from Figure 2 that the groups differed from each other at each successive stage of practice. The no-knowledge group was generally the worst. The results of single classification analyses of variance summarized in Table 3 attest to the over-all significance of these differences for each set of stimuli. All the knowledge groups became significantly superior to the no-knowledge group by the end of practice, as shown by t-tests summarized in Table 4.

TABLE 3

Significance of Differences among All Groups in Mean Error per Stimulus for Sets of Stimuli (F)

Error Measure	Stimulus Set I	II	III	IV	V
Absolute	2.99* (4634)	4.00** (3216)	4.87** h(4104)	5.62** h(3686)	5.70** h(3573)
Algebraic	0.89 h(12744)	0.88 (10151)	1.72 (13045)	2.83* (10523)	3.00* (10275)

* $p < .05$ ** $p < .01$ (5 and 84 df)
h - variance heterogeneous according to Cochran's test

TABLE 4

Significance of Differences in Mean Error Per Stimulus between the No-knowledge and Each of the Knowledge Groups for Sets of Stimuli (t)

Knowledge Group Compared	Absolute Error I	II	III	IV	V	Algebraic Error IV	V
II	1.19 (28)	0.64 (28)	1.14 (28)	0.88 (28)	2.07* (28)	0.88 (25)	0.58 (22)
III	1.62 (28)	0.67 (28)	1.60 (28)	2.11* (28)	2.35* (23)	2.23* (24)	1.69 (24)
IV	2.86** (28)	0.98 (28)	2.85** (24)	3.02** (21)	2.84** (21)	2.07 (17)	6.57** (18)
V	4.01** (28)	3.66** (21)	4.28** (17)	3.77** (16)	4.06** (15)	2.16* (17)	2.33* (13)
VI	3.97** (28)	2.38* (28)	3.92** (21)	3.29** (23)	3.54** (18)	1.99 (15)	1.99 (16)

* $p < .05$ ** $p < .01$ (df given in parentheses)

Figure 3 displays mean algebraic error per stimulus plotted against sets of stimuli. (See Appendix D for tabled values.) Mean algebraic error on the first stimulus was found to be -0.7, 0.6, -1.7, -1.4,-0.8, and -1.0 for Groups I through VI, respectively. These means do not differ significantly (F-ratio of 0.08) and so the groups are regarded as initially equal in algebraic error. In contrast with absolute error, most of the groups made slight constant error initially and maintained their level subsequently. Groups I and II are apparent exceptions in that they tended to progressive overestimation with practice. However, the differences between first-set and last-set performance are in no case reliable (cf. Table 2). Variability around mean algebraic error was great, as is indicated by the standard deviations given in Table 5. While mean standard deviation per stimulus tended to decline with practice, only in the case of Group VI is the difference between variance on the first and last set of stimuli reliable. (For Groups I through VI, respectively, F-ratios are 1.29, 2.12, 2.29, 1.24, 2.10, and 4.42.)

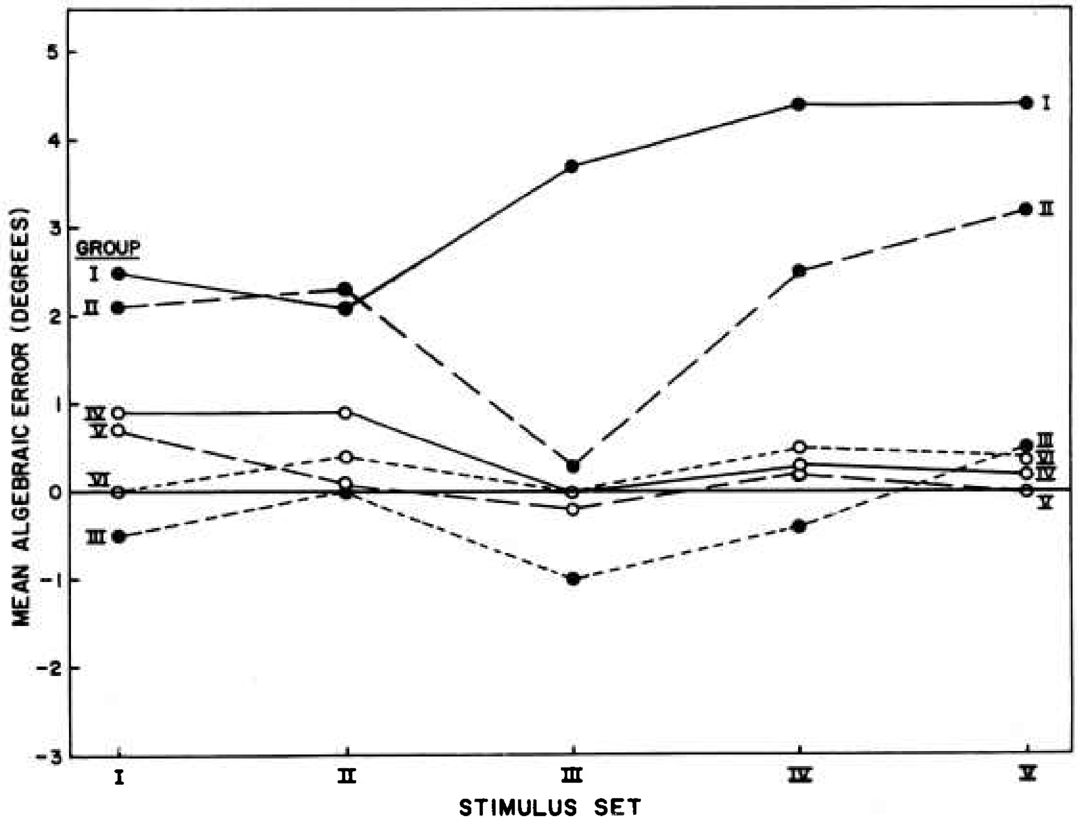


Figure 3. Mean Algebraic Error Per Stimulus Plotted Against Sets of Stimuli.

TABLE 5

Mean Standard Deviation Per Stimulus
for Selected Blocks of Stimuli (s)

Group	I, 1-6	III, 19-24	IV, 1-6	V, 19-24
	Set and Stimuli			
I	11.81	10.67	9.84	9.93
II	11.50	9.20	9.33	7.30
III	9.72	7.98	7.93	6.52
IV	10.46	7.76	5.97	8.17
V	8.63	5.79	5.06	5.55
VI	8.41	6.32	6.43	6.29

The groups do appear to differ in their algebraic error even though they are not so well-ordered by this measure. Over-all tests of the differences among all groups show them nonsignificant for the first three stimulus sets. However, the F-ratios for the last two sets are significant, as is shown in Table 3. Fewer than half the differences between the no-knowledge group's performance and each of the knowledge group's performance on the last two sets are significant (cf. Table 4). The no-knowledge group overestimated significantly more than Groups III and V in the fourth set and Groups IV and V in the last set.

Effects of Specificity. Figure 2 makes it apparent that learning, as measured by absolute error, was generally greater the more specific the knowledge of results given. With but two exceptions (Groups I and II on the second set and Groups V and VI throughout) the groups are perfectly ordered according to specificity. Actually, since Groups II and III received the same kind of information, the difference in their treatment is not properly described as a difference in specificity. Instead, these groups differed in the precision or accuracy of the specific information given them. In consideration of this and the lack of significant difference between them (cf. Appendix E for *t*-values), Groups II and III were combined for the purpose of testing the over-all effect on absolute error of variations in specificity. As summarized in Table 6, the analyses of variance comparing the knowledge groups only show the over-all effects of specificity significant for each stimulus set except the first.

The significance of differences in mean absolute error between pairs of knowledge groups was determined by applying the *t*-test, with degrees of freedom adjusted for variance differences. In general, adjacent treatment groups did not differ significantly, although more extreme ones did. Thus Group V was significantly superior to the combined Groups II and III on all stimulus sets, as was Group VI after the first set. Group V and Group VI never differed significantly, nor was either one of these groups ever significantly superior to Group IV. Group IV was reliably superior to the

combined Groups II and III on the fourth set only. The order of significance and the values of t closely parallel the ordering of groups according to specificity, with Groups V and VI juxtaposed. Appendix E contains a summary of these analyses.

TABLE 6

Significance of Differences among Knowledge Groups (II and III combined) in Mean Absolute Error Per Stimulus for Sets of Stimuli (F)

	Stimulus Set				
	I	II	III	IV	V
F-ratio	1.87	5.17**	4.04*	4.15**	3.49*
Within Groups Var. Estimate	4758	3152	(h)3515	2827	2270

* $p < .05$ ** $p < .01$ (3 and 71 df)
h - variance heterogeneous according to Cochran's test

The performance of the various knowledge groups, as measured by algebraic error and illustrated by Figure 3, was not nearly so well-ordered according to specificity. Only Group II appears to have made much constant error, having displayed a pronounced tendency to overestimate. However, Group II was not significantly inferior to Group III on either the fourth or the fifth set (t values of 1.81 and 1.69 respectively, with 28 degrees of freedom). Since Group III and the other knowledge groups displayed quite similar performance and Group II only was deviant, it did not seem justifiable to combine Groups II and III for the purpose of an over-all test. Instead, Groups II and III were separately contrasted with Groups IV, V, and VI in analyses of variance of performance on the fourth and fifth sets. Nonsignificant F-ratios of 0.37 and 0.09, respectively, were obtained using Group III (within groups estimates of 4448 and 4918, 3 and 56 df). The parallel analysis using Group II resulted in a nonsignificant F of 2.12 for the fourth set and a significant F of 4.90 for the fifth set (within groups estimates of 4823 and 4145). The variances of the groups compared are heterogeneous according to Cochran's test in all four cases. Follow up t-tests showed Group II significantly inferior to Groups IV, V, and VI on the fifth set (t equal 2.44, 2.92, and 2.51, with 28, 19, and 20 df for the respective comparisons).

DISCUSSION AND CONCLUSIONS

Effects of Knowledge. The first hypothesis under test seems clearly confirmed by the results obtained. Rate and level of learning to estimate angular separation are increased when knowledge of results is given. Taken as a group, those receiving no knowledge cannot be said to have learned anything because they showed no demonstrable change in either absolute or algebraic error. However, consideration of their initial performance on the very first stimulus and their variability throughout suggests the possibility that more powerful measures would have shown them to have progressed to an increasingly consistent level of overestimation. In contrast, those receiving knowledge of results in all cases showed significant improvement in absolute error and generally reliable superiority to those receiving no knowledge, particularly in the last stages of practice.

A cursory analysis of the various groups' performance with respect to difficult and easy stimuli points up the effect of knowledge very nicely. The six hardest and the six easiest items on first presentation to the no-knowledge group were determined on the basis of mean absolute error. Then mean absolute error per stimulus was computed separately for each group's performance on the hard and easy items in the first and last stimulus sets. Finally, the percentage by which error was reduced in the last set over the first set was computed. All groups reduced their absolute error in later performance on the hard items by approximately 30% (respectively 30, 29, 31, 34, 26, and 39%). But, on easy items the group receiving no knowledge increased its absolute error 70%, while knowledge groups either stayed at the same level or decreased (respectively 9, -2, 32, 22, and 19%). (As a brief check on the possibility this effect resulted from the manner of selecting hard and easy items, the same percentage was computed for Group VI on items which it found easiest on first encounter. The result was 22%.) Thus, while practice with or without knowledge permitted a reduction in absolute error on the hardest stimuli, practice with knowledge was necessary for the maintenance and improvement of performance on the easiest items. Several other comparisons relating individual differences, stimuli, treatments, and performance to each other might profitably be made.

Effects of Specificity. The second hypothesis also is regarded confirmed by the results obtained. Rate and level of learning to estimate angular separation are directly related to the specificity of the knowledge of results given. The groups were found to be almost perfectly ordered in mean absolute error according to specificity and the significance of this ordering was substantiated by F-ratios in all except the first stimulus set. In addition, paired comparisons of the group means showed the more extreme groups to differ significantly. The failure to find reliable differences among adjacent treatment groups is readily understandable in the light of the considerable variability shown by even the best group. A review of the particular ways in which performance changed in each of the groups points up the effects of specificity.

All the knowledge groups were reasonably equal initially in both absolute and algebraic error. With practice, all showed some unconfirmed tendency to make more consistent responses and all showed a reliable reduction in absolute error. Otherwise performance changes with practice differed. Individuals receiving merely a knowledge of the correctness of their responses (Groups II and III) reduced their absolute error very gradually and never did reach the level of performance of those receiving additional information. Also, while they displayed less constant error than those receiving no information, they appeared worse in this respect than individuals receiving the more specific information. Since the difference in knowledge given to Groups II and III is one of accuracy or precision of information rather than of specificity, any differences in their performance should not be attributed to differences in specificity. Thus, although they did not differ reliably, those encountering the narrower scoring band (III) tended to show less of both types of error. Interestingly, the narrower scoring band seemed to result in a tendency to underestimate, while the wider one produced a confirmed tendency to overestimate. Since the groups differ in this respect, it seems likely that the significant overestimation of Group II was a function of the lack of precise information rather than of the lack of specific information.

Individuals receiving information as to the direction of their error, as well as knowledge of whether they made one (Group IV), reduced their absolute error more rapidly and also performed with very little constant error. While their performance was at no time significantly superior to that of the narrow-tolerance-range group, it was, on occasion, significantly superior to that of the wide-tolerance-range group. Thus, it can be argued that under some combinations of specificity and precision of information, directional information is superior to simple knowledge of correctness of response. Directional information did not result in less apparent variability around constant error.

Ss of both Groups V and VI received relatively specific information about their responses. In addition to knowing whether they had erred and, if so, in what direction, they also knew by what amount. As an apparent result, they reduced their absolute error quite rapidly and to the lowest level of all, were the most consistent, and made very little constant error. While these two groups did not differ significantly, Group V consistently showed the least absolute error. This tendency may be explained in at least two ways. It may be argued that Group V was simply superior to Group VI by virtue of individual differences, as is suggested by their respective errors on the first stimulus. Alternatively, it may be argued that Group VI Ss were slightly inferior because they were forced to perform a subtraction before they could make best use of their knowledge of results. They may have had to find the difference between their response and the correct one and note the direction before they could adequately use the information to guide future responses. If that is true, then it reasonably may be contended that to give knowledge of direction and amount of error is to provide more specific information than that afforded by simply a knowledge of the correct response. A further, more sensitive comparison of the effects of these two types of knowledge obviously would be of interest.

In summary, the effects reviewed are generally substantiated by statistical tests. It is of additional interest that they are exactly what might be expected on common-sense grounds. Information as to whether he had made a wrong response would help an individual to reduce his error somewhat, but only to the extent he could guess which way and by how much he had deviated. As a consequence, he would be prone to considerable variability and tendency to drift into either over- or underestimation. Additional information as to direction of error would help him avoid constant error and, as a consequence, reduce his absolute error without necessarily changing his basic variability very much. Still further information as to amount of error would enable him to perform more consistently and, in so doing, decrease his absolute error as well.

Normative Aspects and Stimulus Effects. Quite aside from providing information about knowledge-of-results conditions, this study also provides useful information on normative performance of angle estimation, both before and after training. As may be seen in Figures 2 and 3 and Table 5, in the absence of information about their performance, Ss quickly tended to overestimate these stimuli by 2 to 3 degrees with very great variability (s in the order of 11 degrees). Their mean absolute error was over 9 degrees. As they continued their estimations, they reduced their variability and mean absolute error a degree or so but increased their tendency to overestimate to about 4.5 degrees. In contrast, those who received training with knowledge of results moved from similar initial performance levels to terminal performance of almost undetectable constant error with half the variability (s approximately 6 degrees) and an absolute error of only around 4 degrees. The performance of the trained Ss seems surprisingly accurate in view of the nature of the stimuli. The performance of the untrained corresponds reasonably well to that of other Ss tested with different materials (Reese, 1953; Baker and Grether, 1954).

The stimuli used in this experiment differed considerably in difficulty. For example, the mean absolute error on individual stimuli of those receiving no information about their responses ranged from 4.5 degrees to 14.5 degrees on first presentation. This lack of consistency in stimulus difficulty reduced the power of the experiment in that it contributed to greater within-group variability.

Although it was not comprehensively attempted, an analysis of the effects of various stimulus parameters on performance could clarify the nature of the perceptual processes involved in angle estimation. Certainly, it would be useful to know what stimulus characteristics are the most important correlates of this particular angle estimation response and how manipulations of them affect the response. Some characteristics of the stimuli used in this experiment were always the same while others varied in a precise way among a limited number of alternatives. Still other characteristics, such as the projection and the perpendicular distance of the small arrow from the command heading line, varied within limits in a haphazard fashion. An analysis was performed with respect to only one stimulus characteristic which was of special interest.

In the course of running Ss one E, Mr. William Pearson, began to suspect that the projection distance (or length of the radius vector as he termed it) played a significant role in determining the accuracy of response. (The projection distance may be defined as the distance from one or the other end of the small arrow along its axis to the command heading line.) As a test of this hypothesis, the projection distance for each stimulus was correlated with mean algebraic error (Pearsonian r) for each group and for the first and last stimulus set or presentation. These correlations and the results of t-tests of their significance seem rather dramatic. On the first presentation, the degree of correlation generally corresponds with the specificity of knowledge ranging from -0.06 to 0.56, with the values of 0.41 and 0.56 for Groups V and VI being the only significant ones. On the last presentation, however, all groups show a reasonably high and significant positive correlation. For Groups I through VI the values obtained are 0.72, 0.58, 0.74, 0.81, 0.76 and 0.82 respectively. This clearly suggests that with practice (with or without knowledge) projection distance emerges as a stimulus aspect to become a very important determiner of the response. There also is the suggestion that knowledge speeds this process.

Generality of Results and Future Research. One final question remains to be considered: To what extent may the principal findings of this study be generalized to other tasks and situations? As was mentioned, the principle of knowledge of results is well supported by research on a variety of tasks. This study merely affords additional evidence for it. However, the generality of the principle that the effectiveness of knowledge of results is a direct function of specificity bears closer scrutiny.

Described in an abstract way, this task required an estimation of spatial extent and a translation of the estimate into a scaled verbal response. Thus, an absolute judgment of spatial extent had to be made. It is a reasonable expectation that parallel variations in specificity would have the same general effects on the acquisition of skill in making other absolute judgments of spatial extent. Of somewhat less certainty, however, is whether the same effects would be observed when a relative judgment or comparison of spatial extent was required or when judgments about other aspects of visual stimuli were required. In still greater doubt is whether the acquisition of skill in making judgments about stimuli presented to other sense modalities would be similarly influenced by variations in specificity. In general, the degree of certainty with which it can be predicted that increased specificity in knowledge of results will speed the acquisition of a perceptual skill depends directly on the similarity between the task involved and the one used in this study.

Somewhat further along the dimension of task similarity are a variety of other tasks not generally regarded as perceptual. The applicability of these findings to such nonperceptual tasks may be even more seriously disputed. Additional research on the effects of variations in the specificity of knowledge on learning other types of tasks would be desirable.

There is another kind of limitation on the generality of these results. They do not permit the conclusion, even for the same task, that skill improved through practice with relatively specific knowledge of results will be strongly retained when knowledge of results is discontinued or that correspondingly great transfer will result in performing a similar task. In short, hypotheses about retention and transfer, paralleling those tested by this experiment, should be tested by further research.

SUMMARY

This experiment was designed to test two related hypotheses regarding the effects of knowledge of results on the learning of a perceptual skill. The hypotheses are that (1) rate and level of learning to make an absolute judgment of spatial extent will be increased with knowledge of results and (2) this effect will be greater the more specific the knowledge.

Ninety male undergraduate subjects were assigned equally and at random to six treatment groups. Each subject estimated individually with respect to 120 stimulus photos how many degrees a $\frac{1}{4}$-inch arrow would have to be turned to exactly parallel an adjacent arrowheaded line drawn across a $3\frac{1}{2}$-inch circle. The stimuli were grouped into five sets, each set containing copies of the same 24 different stimuli in a different random order. Correct answers ranged from 11 through 44 degrees and were never repeated within a set. The stimuli were presented for 5 seconds with 10 seconds between. Intervals of 2 minutes separated each set of 24, except that 6 minutes separated the third and fourth sets. Knowledge of results given orally after each estimation ranged in specificity from simple right-wrong information to amount and direction of error information for five experimental groups. No knowledge was given a sixth, control group.

Differences in mean absolute and algebraic error per stimulus among groups reflect differences in the knowledge given. With practice, those who were given <u>no information</u> showed no demonstrable change in performance. They made large absolute error (over 8°) and tended toward increasingly consistent and significant overestimation (about 4.5° with s about 11°). Those who were either given the <u>correct answer</u> or told by what <u>amount</u> and in what <u>direction</u> they erred showed almost undetectable algebraic error with half the variability (s about 6°) and very rapid reduction of absolute error to the lowest level (about 4°). Between these extremes was the performance of those who received <u>right-wrong</u> information (with reference to two tolerance ranges) and those who received information as to <u>direction</u> of error only. Directional information was superior to right-wrong information (or simple knowledge of error) as shown by both absolute and algebraic error.

In general, all five knowledge groups showed significant improvement with practice and significant superiority to the no-knowledge group in absolute error. In addition, the knowledge groups were almost perfectly ordered in absolute error according to the specificity of knowledge given

them. Over-all differences in their performance are significant for all except the first stimulus set, as are the differences between pairs of groups receiving the more extreme treatments. Most knowledge groups made little constant or algebraic error but those who received only the simple knowledge of error with a wide tolerance range displayed significant overestimation. Therefore, both hypotheses are regarded confirmed. However, these findings should not be applied indiscriminantly to tasks not requiring an absolute judgment of spatial extent or to problems of transfer and retention.

REFERENCES

Baker, C. A. and Grether, W. F. Visual presentation of information. USAF WADC Tech. Rep., 1954, No. 54-160.

Brown, J. S. A proposed program of research on psychological feedback (knowledge of results) in the performance of psychomotor tasks. In R. M. Gagne (Ed.) Research Planning Conference on Perceptual and Motor Skills. USAF HRRC Conf. Rep., 1949, No. 49-2. (pp. 81-87)

Eriksen, C. W. Effects of practice with and without correction upon discrimination learning under absolute conditions. USAF WADC Tech. Rep., 1957, No. 57-71.

Gibson, Eleanor J. Improvement in perceptual judgments as a function of controlled practice or training. Psychol. Bull., 1953, 50, 401-431.

Gilliland, A. R. The effect of practice with and without knowledge of results in grading handwriting. J. educ. Psychol., 1925, 16, 532-536.

Goldstein, M. and Rittenhouse, C. H. Knowledge of results in the acquisition and transfer of a gunnery skill. J. exp. Psychol., 1954, 48, 187-196.

Hamilton, H. C. The effect of incentives on accuracy of discrimination measured on the Galton bar. Arch. Psychol., 1929, 16, No. 103.

Hirsch, R. S. The effects of knowledge of test results on learning of meaningful material. USN Special Devices Cent. Hum. Engr. Rep., 1952, No. 269-7-30.

Minturn, A. and Reese, T. W. The effect of differential reinforcement on the discrimination of visual number. J. Psychol., 1951, 31, 201-231.

Parks, R. B. Verbal feedback in pedestal sight manipulation. USAF PTRC Res. Bull., 1954, No. 54-91.

Reese, E. P. (Ed.) Psychophysical research. Summary report 1946-1952. USN Special Devices Cent. Tech. Rep., 1953, No. 131-1-5.

Rock, R. T., Jr. Knowledge of results of responses compared with after-effects following each response in a learning situation. Psychol. Bull., 1936, 33, 613. (abstract)

Ross, C. C. An experiment in motivation. J. educ. Psychol., 1927, 18, 337-346.

Ross, C. C. The influence upon achievement of a knowledge of progress. J. educ. Psychol., 1933, 24, 609-619.

Solomons, L. Discrimination in cutaneous sensations. Psychol. Rev., 1897, 4, 246-250.

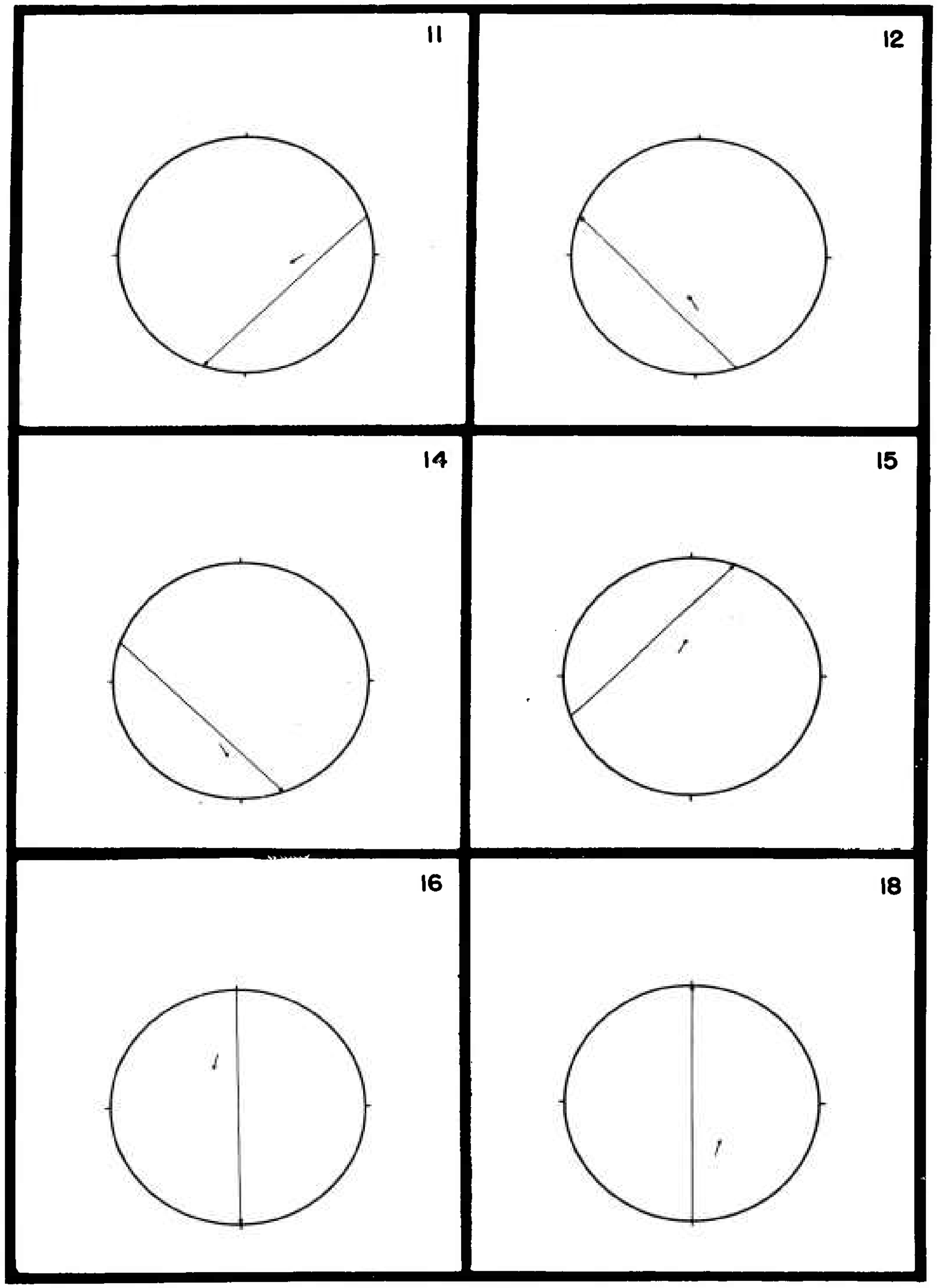
Taubman, R. E. The effect of practice with and without reinforcement of the judgment of auditory number. J. exp. Psychol., 1944, 34, 143-151.

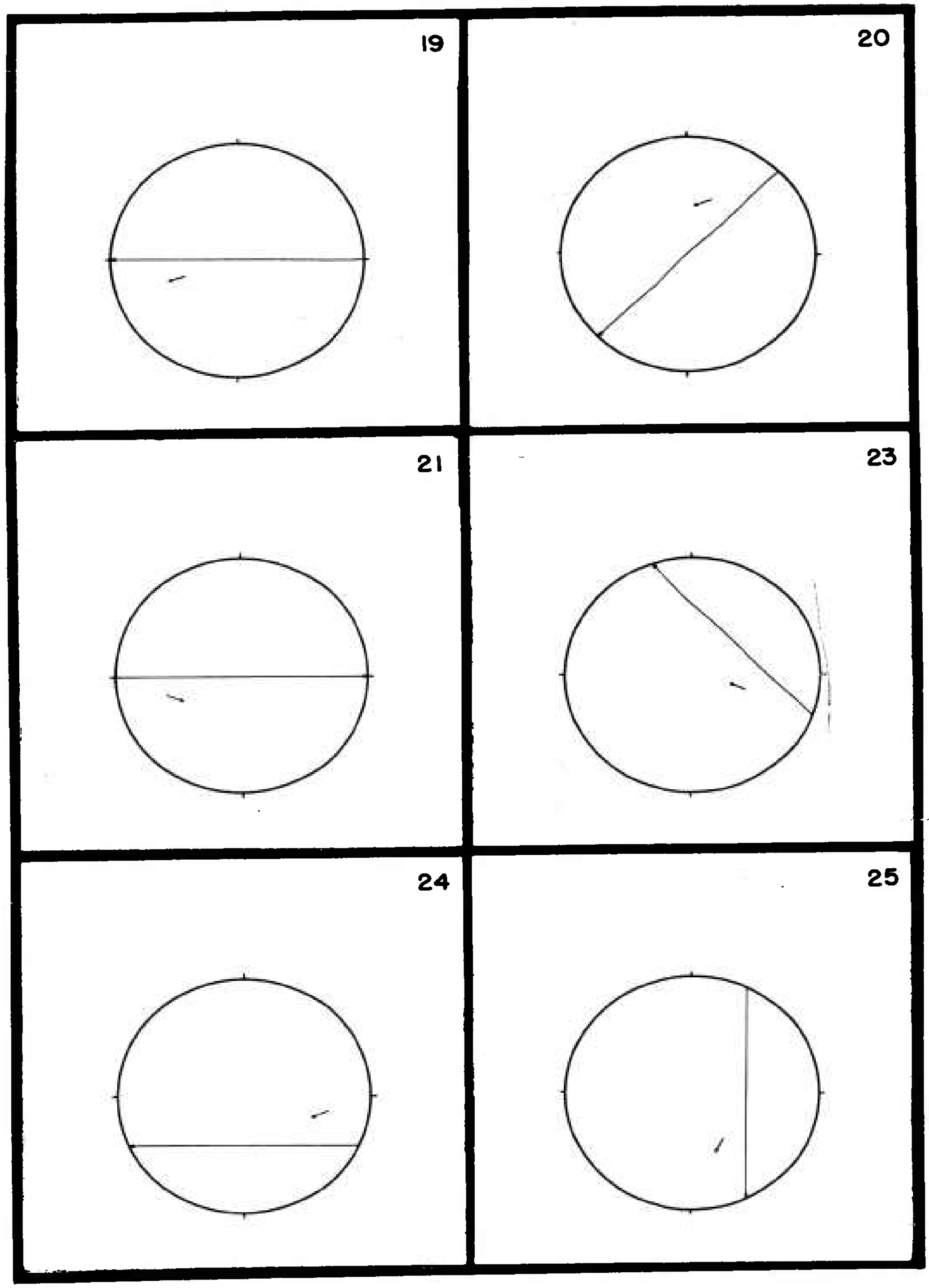
Thorndike, E. L. The law of effect. Amer. J. Psychol., 1927, 39, 212-222.

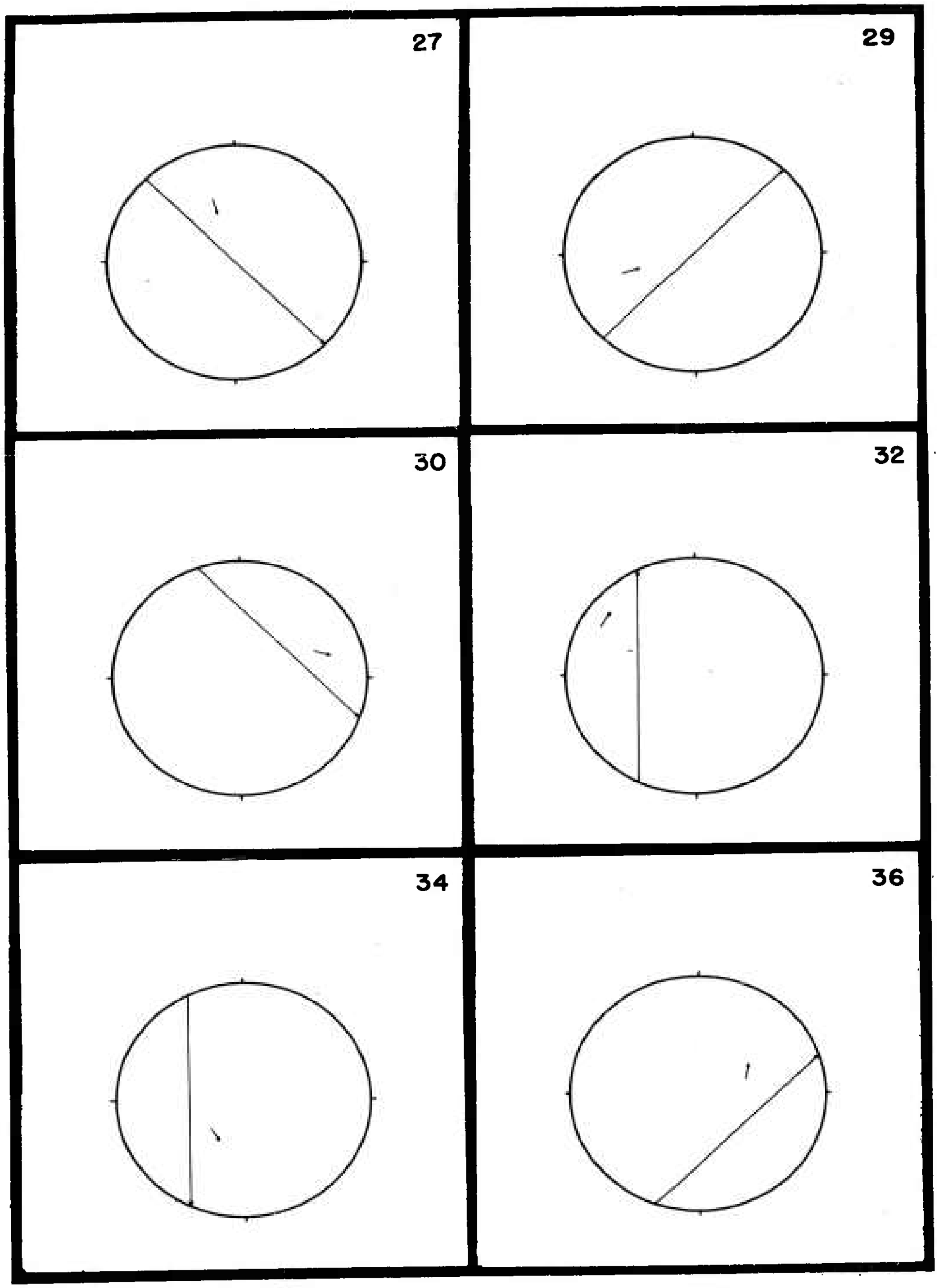
Trowbridge, M. H. and Cason, H. An experimental study of Thorndike's theory of learning. J. gen. Psychol., 1932, 7, 245-260.

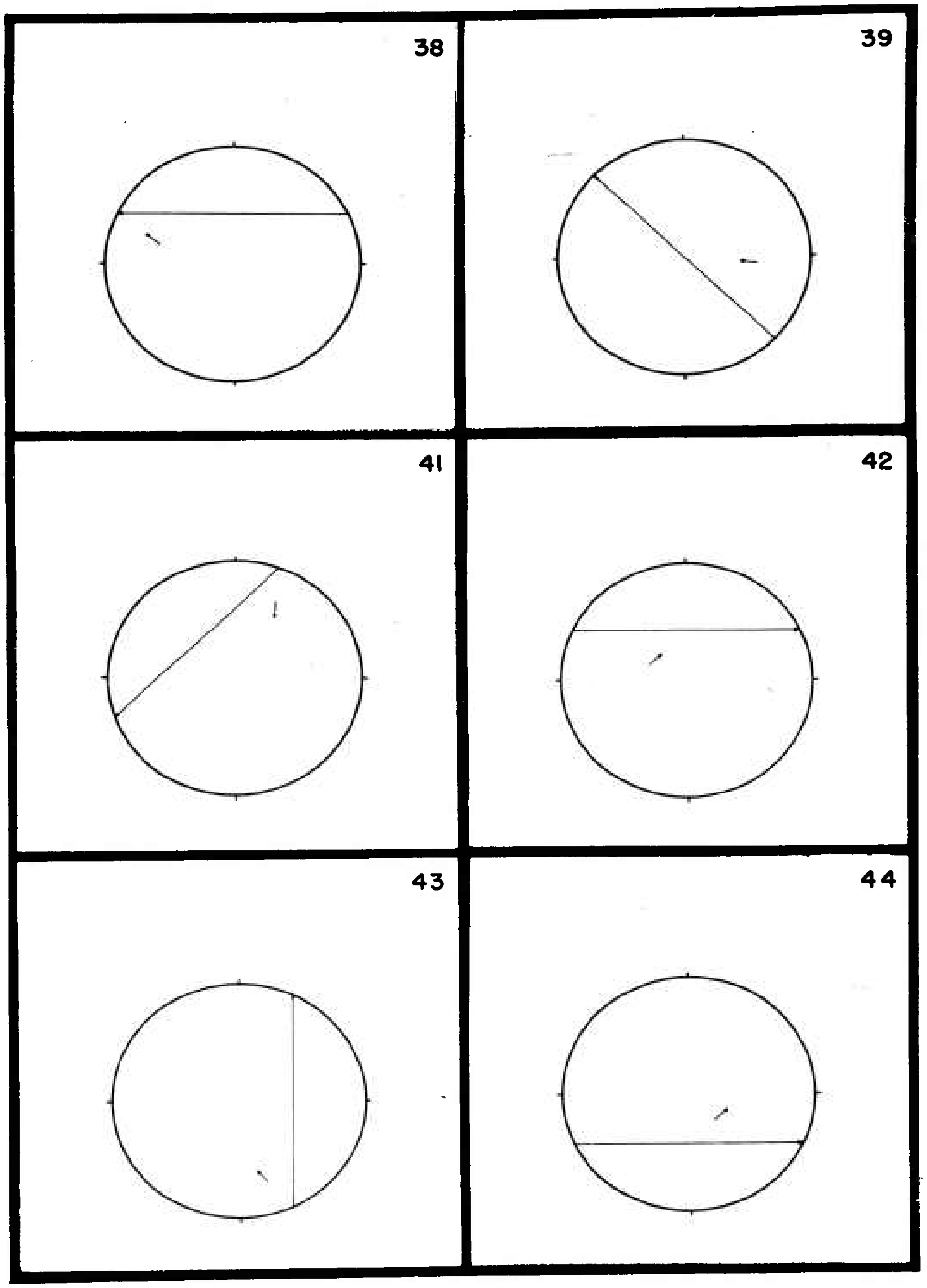
Waters, R. H. The specificity of knowledge of results and improvement. Psychol. Bull., 1933, 30, 673. (abstract)

APPENDIX A. The Stimuli (half-size with correct answers).




19
20
21
23
24
25


27
29
30
32
34
36


38
39
41
42
43
44

APPENDIX B

Instructions to Subjects

"Are you familiar with how angles are measured? Well, if you think of a right angle like the corner of a square, that would be 90 degrees. Half of that would be 45 degrees, and so forth.

"In this experiment I'm going to show you a series of stimulus cards much like this one. (E holds up sample.) You'll see them each time through an opening here (E points). For each card imagine that you are flying a plane in the direction shown by this small arrow. The long line with the arrow on one end shows the direction in which you are supposed to fly. I'd like you to tell me how many degrees you would have to turn to fly in the same direction as the long line, that is, so that the small arrow would be parallel to the line.

"Do this each time immediately after I close the opening.

"I'll let you look at each one for about 5 seconds, wait 10 seconds, show you another for 5 seconds and so forth until we have completed the series. Then I'll wait 2 minutes before beginning the next one.

"Remember to give your answer each time without hesitation or delay as soon as I close the opening.

Group I: (E gives no further instructions.)

Groups II and III: "Shortly afterward I will say 'right' or 'wrong' to show you whether you have made an error.

Group IV: "Shortly afterward I will say 'over,' 'under,' or 'correct,' to show you the direction of your error. 'Over' would mean your estimate was too high.

Group V: "Shortly afterward I will say 'over so many degrees' or 'under so many degrees' to show the direction and amount of your error. 'Over two' would mean your estimate was 2 degrees too high. If you are exactly right, I'll say 'correct.'

Group VI: "Shortly afterward I will tell you the correct answer so you will know the direction and amount of your error.

"Do you have any questions?

"Ready?" . . .

APPENDIX C

Mean Absolute Error Per Stimulus in Degrees

Set and Stimuli	Group I	II	III	IV	V	VI
I, 1-6	9.4	8.5	7.3	7.3	6.7	6.5
I, 7-12	10.5	8.0	8.7	6.7	6.7	6.2
I, 13-18	8.3	8.1	7.8	6.6	5.6	6.2
I, 19-24	9.5	8.0	7.1	7.0	5.3	6.5
I, All	9.4	8.2	7.7	6.9	6.1	6.3
II, 1-6	6.3	7.1	5.2	6.3	4.3	5.0
II, 7-12	8.5	8.5	7.2	6.7	4.4	5.8
II, 13-18	6.8	7.8	6.8	5.3	4.7	5.5
II, 19-24	7.0	7.7	7.0	6.4	4.5	4.6
II, All	7.2	7.8	6.6	6.2	4.5	5.2
III, 1-6	9.0	7.2	8.2	4.8	4.4	4.5
III, 7-12	8.0	6.6	6.4	6.1	4.6	4.8
III, 13-18	8.3	7.3	5.5	5.9	4.4	4.3
III, 19-24	8.6	6.8	6.4	5.7	4.7	5.0
III, All	8.5	7.0	6.6	5.6	4.5	4.7
IV, 1-6	8.6	7.5	6.2	4.9	3.8	5.3
IV, 7-12	8.3	8.1	5.5	4.4	4.5	4.4
IV, 13-18	8.8	7.2	4.9	6.0	5.0	4.7
IV, 19-24	7.9	6.7	7.0	5.2	5.0	4.4
IV, All	8.4	7.4	5.9	5.1	4.6	4.7
V, 1-6	8.8	5.9	5.9	4.8	3.5	3.9
V, 7-12	8.8	7.8	5.6	4.9	4.4	4.5
V, 13-18	7.9	5.1	6.6	5.3	4.0	4.9
V, 19-24	8.2	5.7	4.8	5.6	4.6	4.6
V, All	8.4	6.1	5.7	5.2	4.1	4.5

APPENDIX D

Mean Algebraic Error Per Stimulus in Degrees

Set and Stimuli	Group I	II	III	IV	V	VI
I, 1-6	3.0	1.9	-1.0	0.6	0.8	-0.5
I, 7-12	2.7	1.1	-1.6	1.9	1.9	1.2
I, 13-18	3.0	3.7	-0.3	0.6	0.2	1.5
I, 19-24	1.2	1.7	0.9	0.4	0.0	-2.1
I, All	2.5	2.1	-0.5	0.9	0.7	0.0
II, 1-6	0.5	2.8	0.7	1.1	0.8	0.6
II, 7-12	3.5	3.7	0.9	0.9	0.3	0.8
II, 13-18	2.7	2.4	0.5	0.7	0.0	1.2
II, 19-24	1.8	0.4	-2.0	0.9	-0.8	-1.1
II, All	2.1	2.3	0.0	0.9	0.1	0.4
III, 1-6	4.4	1.4	-0.6	0.6	-0.7	0.5
III, 7-12	4.7	-0.7	-1.0	0.9	0.9	0.3
III, 13-18	3.3	0.8	-0.6	-0.8	-0.2	0.1
III, 19-24	2.3	-0.3	-1.4	-0.5	-0.7	-0.1
III, All	3.7	0.3	-1.0	0.0	-0.2	0.0
IV, 1-6	4.8	1.9	-0.5	0.7	0.9	1.7
IV, 7-12	4.4	2.1	0.7	1.4	-0.6	0.5
IV, 13-18	5.1	3.9	-1.0	0.2	0.2	0.4
IV, 19-24	3.5	2.1	-1.1	-1.2	0.2	-0.7
IV, All	4.4	2.5	-0.4	0.3	0.2	0.5
V, 1-6	5.2	3.4	0.8	0.1	-0.2	0.1
V, 7-12	4.8	4.2	0.1	0.3	0.0	1.2
V, 13-18	4.2	2.7	0.7	-1.2	0.2	0.2
V, 19-24	3.4	2.5	0.4	1.4	-0.2	0.0
V, All	4.4	3.2	0.5	0.2	0.0	0.4

APPENDIX E

Significance of Differences in Mean Absolute Error Per Stimulus between Pairs of Knowledge Groups for Sets of Stimuli (t)

Knowledge Groups Compared	Stimulus Set				
	I	II	III	IV	V
II vs. III	0.38 (28)	1.31 (28)	0.27 (28)	1.15 (28)	0.18 (28)
II vs. IV	1.15 (28)	1.54 (28)	1.28 (23)	2.21* (24)	0.75 (28)
II vs. V	1.94 (28)	4.14** (19)	2.41* (16)	3.18** (19)	2.24* (15)
II vs. VI	1.74 (22)	2.94** (28)	2.20* (20)	2.59* (28)	1.61 (23)
III vs. IV	0.71 (28)	0.39 (28)	1.14 (28)	1.06 (28)	0.65 (28)
III vs. V	1.48 (28)	3.05** (22)	2.61* (18)	2.12* (21)	2.53* (16)
III vs. VI	1.25 (22)	1.74 (28)	2.37* (24)	1.56 (28)	1.67 (28)
II and III vs. IV	1.10 (43)	1.11 (43)	1.52 (43)	2.15* (42)	1.12 (43)
II and III vs. V	2.15* (43)	4.71** (43)	3.40** (40)	3.57** (41)	3.61** (40)
II and III vs. VI	1.95 (43)	2.88** (43)	2.90** (43)	2.54* (43)	2.39* (43)
IV vs. V	0.90 (28)	2.05 (19)	1.78 (22)	1.22 (28)	1.90 (17)
IV vs. VI	0.58 (28)	1.07 (25)	1.38 (28)	0.69 (28)	1.04 (28)
V vs. VI	0.42 (28)	1.25 (25)	0.02 (28)	0.26 (22)	0.78 (18)

* $p < .05$ ** $p < .01$ (df given in parentheses)